Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 69 (2012) 2137 – 2142

Procedia
Social and Behavioral Sciences

International Conference on Education and Educational Psychology (ICEEPSY 2012)

An Analysis of the Teachers' Sense of Efficacy Scale within the Malaysian Context using the Rasch Measurement Model

Ahmad Zamri bin Khairani[a*], Nordin bin Abd. Razak[b]

[a,b]*School of Educational Studies, 11800 Universiti Sains Malaysia*

Abstract

Many concerns have been raised about the psychometric properties of instruments when used in different culture. Studies show that adaptation of instrument within a particular context sometimes has different meaning since a particular psychology construct is considered cultural specific. The purpose of the present study is to examine the psychometric properties of the Teachers' Sense of Efficacy Scale (TSES) within the Malaysian educational context A total of 191 Malaysian in-service and 122 pre-service teachers participated in the study. A Rasch Measurement Model analysis is used to examine the psychometric properties of the scale, particularly the reliability and construct validity. Analyses showed somewhat similar variability with the in-service group demonstrated wider spread of item difficulty compared to the pre-service teachers. One important observation was that both group endorsed different set of items. Consistency of both item difficulty and teachers' efficacy were high and threats to construct validity in terms of construct irrelevant-variance and construct under-representation were minimum. One worrying finding, however, was that the differential item function (DIF) analysis showed that 13 items (54.17%) behaved significantly different across groups of teachers. Even though several items shown to behave differently across both group of teachers, overall results are encouraging and seem to support the suitability of the TSES to assess teachers' sense of efficacy within the Malaysian educational context

Selection and peer-review under responsibility of Dr. Zafer Bekirogullari of Cognitive – Counselling, Research & Conference Services C-crcs.

Keywords: In-service teachers; pre-service teachers; teaching efficacy; Rasch Measurement Model

Selection and peer-review under responsibility of Dr. Zafer Bekirogullari of Cognitive – Counselling, Research & Conference Services C-crcs.
doi:10.1016/j.sbspro.2012.12.178

1. Introduction

Teacher's sense of efficacy is defined as one's belief of his or her capabilities to bring about desired outcomes of students' engagement', achievement and learning, even among those students who may be difficult or unmotivated Bandura (1997). The construct has been the focus of many studies due to the fact that the construct has established itself as having significant implications especially in educational settings. Strong sense of teaching efficacy is related to teachers' positive behaviours, which in turn, enhance students' achievement. Highly efficacious teachers are more open to new ideas, more committed and invest greater effort in teaching, as well as provide more attention to low-achieving students (Hoy & Spero, 2005; Tschannen-Moran & Woolfolk Hoy, 2001). Teacher's sense of efficacy has also been associated with various positive aspects in teaching profession such as (1) more enthusiastic and satisfied with their job (Caprara, Barbaranelli, Steca, & Malone, 2006), and (2) demonstrate lower levels of burnout (Friedman, 2003; Schwarzer & Hallum, 2008). Because of this, Caprara, Barbaranelli, Borgogni, and Steca, (2003) speculate that highly efficacious teachers will be successful in teaching as well as able to function well in school environment.

Luszczynska, Scholz, and Schwarzer (2005) observe that over the years, a considerable amount of instruments have been designed to assess teachers' sense of efficacy. Nevertheless, a number of researchers have raised concerns about the adequacy of the instrument used, particularly the validity of both the measure as well as the interpretation (Henson, 2002). In addition, Tschannen-Moran and Woolfolk Hoy (2001) show that the existing instruments also demonstrates both conceptual and methodological shortcomings. In an effort to overcome the shortcomings, they develop the TSES that comprises 24 items measuring three sub dimensions of teachers' sense of efficacy, namely, instructional strategies, classroom management, and student engagement. Instructional strategies measures the extent teachers can play their part as classroom instructors to facilitate both conducive learning environment and effective learning process. This involves teaching strategies, teaching approach and handling of teaching and learning and factor. Efficacy for classroom management, meanwhile, measures to what extent teachers can manage students discipline and control students' behaviour. Meanwhile, efficacy for student engagement, measures to what extent teachers believe they can install all the positive attitudes among their students such as influencing students to get engaged in school activities, increasing students' motivation and others.

Hoy and Spero (2005) find that the TSES scale exhibits a stable factor structure that is an integral part for a quality measurement. Thus, it is not unexpected that the TSES has been widely employed in various setting and samples such as both in-service and pre-service teachers (Tschannen-Moran & Woolfolk Hoy, 2001; Larson, & Goebel, 2008), early childhood educators (Brown, 2005), and special education teachers (Smith, 2008). Nevertheless, being a new instrument, the TSES still lacks empirical data concerning its psychometric properties especially in terms of cross cultural context. Studies show that adaptation of teaching efficacy instrument within a particular context is inconclusive. Lin and Gorell (2001), for example, have been able to demonstrate that the construct is cultural specific based on their study in Taiwan. In contrast, Tsigilis, Koustelios, and Grammatikopoulos, (2010) find that administration of TSES to a sample of Greek educators did not influence its psychometric properties. Since the TSES is one of the most widely used scale within the Malaysian context (Abdul Rahim, Mohd Majid, Rashid, & Lyndon, 2008; Khalid, Zurida, Shuki & Ahmad, 2009), it is about time a study is conducted to provide better understanding of its' psychometric properties

2. Method

2.1 Participants

A convenience sample of 191 in-service teachers and 122 pre-service teachers participated in the present study. The in-service teachers comprise of 45 males (23.6%) and 146 females (76.4%) from local public schools. The pre-service teachers, sampled from a local public university, the ratio between male and female stands at 21 (17.2%) to 101 (82.8%). With regards to the in-service teachers, their mean for teaching experience is 9.74 years (SD = 3.85 years). Inclusion of both groups of teachers is replicated from the method employed by Tschannen-Moran and Woolfolk Hoy (2001) during development of the TSES.

2.2 The instrument

The TSES consists of 3 sub dimensions of teachers' sense of efficacy scale. Table 1 provides example of items for every scale.

Table 1: The Teachers' Sense of Efficacy Scale

Sub Dimension	No of Items	Example of Items
Instructional strategies	8	How well can you respond to difficult questions from your students?
Classroom management	8	How much can you do to control disruptive behavior in the classroom?
Student engagement	8	How much can you do to get through to the most difficult students?
Total	24	

2.3 Data Analysis

In order to provide meaning to the score, Rasch Measurement Model analysis using WINSTEPS version 3.57 was employed in this study. The analysis is a method of obtaining objective, fundamental and linear measures from observation of ordered category such as Likert-scale. The Rasch procedure transforms the summated test score into interval-scale score called 'measure' that is calibrated in log-odd or *logits* unit. Rasch analysis offers mathematical framework to evaluate the extent to which the data represents a single construct as well as the extent to which the data fit model expectation. In addition, it also offers appropriate conclusion and assignment of meanings on the scores of the construct. Several appealing outcomes from the procedure include (1) instrument-free respondent measures, (2) respondent-free item difficulty measures, (3) evidence of construct validity of the measures. Moreover, the analysis allows direct comparison between the performances of the two parameters, namely, teachers' sense of efficacy and difficulties of the relevant items along a single continuum (Wright & Masters, 1982). However, consistent with the purpose of the present study, only the latter will be discussed

In order for such a measure to have the abovementioned properties, two important assumptions must be met. Firstly, Rasch Model requires that the data must fit the model, that is, the degree of discrepancy between observed by the data and the expected by the model is kept to a reasonable level (Wright & Masters, 1982). The infit mean square (MNSQ) and outfit MNSQ provide indications of the discrepancies. This study adopts the range of acceptable fit between 0.7 – 1.3 for both fit indices as suggested by Bond and Fox (2001). Secondly, the data must meet the unidimensionality assumption, that is, they represent a single construct (Wright & Masters, 1982). In WINSTEPS, the principal component analysis of the residuals procedure helps identify the existence of second factor that pose a threat to unidimensionality assumption.

The data were analyzed firstly with the whole sample (N = 313) and later the analyses were repeated with each of the two groups of the sample (N = 191 in-service, N = 122 pre-service). On every analysis, descriptive statistics (items means, standards deviations and fit indices) were estimated. Consistency of both ordering of items and person are also reported. Then, the measures of both groups were compared to identify items that have significantly different meanings across groups in a procedure called differential item functioning (DIF) [see Bond and Fox (2001)]. Another analysis conducted is related to construct validity of the measures meaning and interpretation. Based on the foundation laid by Messick (1993) two major threats to construct validity that are under investigation are construct-irrelevant variance and construct under-representation. The former relates to the irrelevant variances that contaminate measurement of the main construct while in the latter, the measurement fails to include important sub-dimensions of the construct. In short, construct validity requires nothing irrelevant be added while at the same time nothing important should be left out in assessing a construct. Within the framework of Rash Measurement Model, Baghaei (2008) suggests that construct-irrelevant variance can be assessed by examining both dimensionality and fit of the measurement while significant gaps between the subsequent items provide indication of construct under-representation

3. Findings and Discussion

Principal component analysis performed on the residuals resulted in the second factor extracted had a strength of about 3 items for all 3 separate analyses. The factor thus did not contain enough information that can pose a threat to the main construct. The Rasch analysis as presented in Table 1 found both means of infit and outfit MNSQ were close to the expected value of 1.00. Inspection with individual items showed that infit MNSQ values ranged from 0.75 to 1.25 while outfit mean-square values ranged from 0.67 to 1.26. The results supported the following: (1) the unidimensionality assumption of the construct was met, and (2) the scores demonstrated little variation from model expectation – that there was evidence of consistency between 313 teachers' response and 24 items on the scale. As depicted in Table 2, all 3 analyses showed somewhat similar variability with the in-service group demonstrated wider spread of item difficulty compared to the pre-service teachers. One important observation was that both group endorsed different set of item. For example, while the in-service teachers endorsed Q8 (*How well can you establish routines to keep activities running smoothly?*) as the easiest item (a lot of agrees and strongly agrees responses), their pre-service counterpart chose Q16 (*How well can you establish a classroom management system with each group of students?*) instead. Similarly, the most difficult-to-endorsed items (a lot of strongly disagrees and disagrees responses) are Q22 (*How much can you assist families in helping their children do well in school*?) and Q1(*How much can you do to get through to the most difficult students?*) respectively for both groups. We might speculate that since they are yet to have experience in schools, the pre-service teachers may have unrealistic views of their teaching efficacy. The mismatch will certainly created problems during their teaching experience later. The in-service teachers demonstrated higher sense of efficacy, but the in-service teachers showed more variability.

The prime analysis in assessing psychometric properties of the TSES is the consistency of the measures and evidence of construct validity. Consistency of the item difficulty was high based on the item reliability coefficients (overall = .94, in-service = .91, pre-service = .87) suggesting that the ordering of item difficulty was replicable with other comparable sample of teachers. Meanwhile, consistency of teachers' efficacy measures was also high, indicating that it was highly likely that the ordering of teacher efficacy can be replicated since most of the variance was attributed to true variance of the teaching efficacy construct. From the findings, threat regarding construct irrelevant-variance was minimum based on the dimensionality test as well as the within-range fit indices. In addition, no gaps of .5 *logits* or more (Linacre, 2004) between subsequent items on the measured scale showed that the TSES was broad enough to include important sub-dimensions of teaching efficacy construct. One worrying finding, however, is that the DIF analysis showed that 13 items (54.17%) behaved significantly different across groups of teachers, in which 8 items favours the in-service teachers while another 5 items favours the pre-service teachers. The item needs to be investigated further since they may indicate bias to a particular group. The analysis, however, is beyond the scope of this present paper.

Table 2: Statistics according to Groups

	Overall				In-service				Pre-service			
			Infit	Outfit			Infit	Outfit			Infit	Outfit
Item	Measure	SE	MNSQ	MNSQ	Measure	SE	MNSQ	MNSQ	Measure	SE	MNSQ	MNSQ
1	0.62	0.11	1.22	1.22	0.42	0.14	1.23	1.24	1.12	0.2	1.27	1.25
2	0.9	0.11	1.23	1.26	1.02	0.14	1.13	1.13	0.77	0.2	1.04	1.05
3	-0.28	0.11	0.95	0.95	-0.6	0.14	0.95	0.93	0.34	0.2	0.91	0.93
4	-0.34	0.12	0.87	0.85	-0.23	0.14	0.91	0.91	-0.61	0.21	0.86	0.79
5	-0.71	0.12	0.99	0.99	-0.6	0.14	1.05	1.06	-1.06	0.22	0.99	0.98
6	-0.73	0.12	0.86	0.87	-0.72	0.15	0.92	0.9	-0.83	0.21	0.86	0.93
7	0.08	0.11	1.14	1.13	-0.25	0.14	1.22	1.22	0.74	0.2	0.97	1.01
8	-0.91	0.12	1.02	1.01	-1.38	0.15	0.91	0.88	-0.1	0.2	1.02	0.97
9	-0.11	0.11	0.93	0.9	-0.11	0.14	0.9	0.89	-0.14	0.2	1.09	0.98
10	-0.28	0.11	0.94	0.92	-0.17	0.14	0.88	0.88	-0.57	0.21	1.2	1.08
11	0.21	0.11	0.85	0.85	0.42	0.14	0.79	0.79	-0.22	0.2	0.98	0.97

12	0.18	0.11	0.97	0.99	0.31	0.14	0.9	0.92	-0.06	0.2	1.18	1.19
13	-0.44	0.12	1.07	1.06	-0.64	0.14	1.17	1.16	-0.1	0.2	0.94	0.88
14	0.14	0.11	0.96	0.97	0.5	0.14	1.01	1.02	-0.61	0.21	0.77	0.69
15	-0.45	0.12	1.06	1.05	-0.81	0.15	1.03	1.04	0.22	0.2	1.01	0.95
16	-0.14	0.11	0.87	0.86	-0.19	0.14	0.97	0.96	-0.06	0.2	0.75	0.67
17	0.47	0.11	1.04	1.04	0.71	0.14	1.07	1.09	0.02	0.2	0.96	0.91
18	0.39	0.11	0.94	0.93	0.59	0.14	0.87	0.85	-0.02	0.2	1.11	1.06
19	0.28	0.11	0.92	0.91	0.09	0.14	0.9	0.89	0.7	0.2	0.99	0.99
20	-0.05	0.11	0.75	0.73	0.03	0.14	0.76	0.76	-0.22	0.2	0.81	0.75
21	0.44	0.11	1.13	1.15	0.19	0.14	1.24	1.24	0.97	0.2	0.95	0.99
22	0.75	0.11	1.21	1.23	1.03	0.14	1.18	1.22	0.48	0.2	0.98	0.88
23	0.04	0.11	0.92	0.9	0.4	0.14	0.89	0.86	-0.7	0.21	0.95	0.94
24	-0.05	0.11	1.08	1.06	-0.03	0.14	1.1	1.08	-0.06	0.2	1.17	1.17
mean	0	0.11	1.00	0.99	0	0.14	1.00	1.00	0	0.2	0.99	0.96
SD	0.47	0	0.13	0.13	0.6	0	0.14	0.15	0.58	0.01	0.13	0.14
Min (item)	-0.91 (Q20)	0	0.13	0.13	-1.38 (Q8)	0	0.14	0.15	-1.06 (Q16)	0.01	0.13	0.14
Max (item)	0.75 (Q2)	0.12	1.21	1.23	1.03 (Q22)	0.15	1.24	1.24	0.97 (Q1)	0.22	1.2	1.19
Range	1.66	0.12	1.08	1.1	2.41	0.15	1.1	1.09	2.03	0.21	1.07	1.05

4. Conclusion

The study reports on the Rasch Measurement Model analysis on teachers' sense of efficacy using the TSES. Its limitation notwithstanding, the present validation study provides some empirical evidence that extends the understanding of teachers 'sense of efficacy construct. The findings showed that the instrument was able to provide valid and reliable estimates of sense of teaching efficacy within the Malaysian contexts. Teachers' perception also fits the requirements of Rasch measurement model, thus providing good psychometric properties of the construct. Nevertheless, the study also showed that some items performed differently across groups. The items need to be analyzed further to investigate determine how serious the discrepancies are.

5. Acknowledgement

This article is made possible by the funding obtained from the Universiti Sains Malaysia

6. References

Abdul Rahim, B., Mohd Majid, K., Rashid, J., & Lyndon, N. (2008). Teaching efficacy of Universiti Putra Malaysia Science student teachers. *College Student Journal, 42*(2), 493-508.

Baghaei, P. (2008). The Rasch Model as a construct validation tool. *Rasch Measurement Transactions*. Retrieved June 21, 2011 from http://www.rasch.org/rmt/ rmt221a.htm.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman (Chapter 1 – 3).

Bond, T. G. & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Brown, E. T. (2005). The influence of teachers' efficacy and beliefs regarding mathematics instruction in the early childhood classroom. *Journal of Early Childhood Teacher Education*, *26*, 239-257.

Hoy, A., and Spero, R.B. (2005). Changes in teacher-efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education,* 21(4), pp. 343-356.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44,* 473-490.

Caprara, G. V., Barbaranelli, C., Borgogni, L., and Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology*, 95, pp. 821-832.

Friedman, I. A. (2003). Self-efficacy and burnout in teaching: The importance of interpersonal-relations efficacy. *Social Psychology of Education*, *6*, 191-215.

Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, *37*, 137-150.

Khalid Johari, and Zurida Ismail, and Shuki Osman,&Ahmad Tajuddin Othman (2009). *Pengaruh jenis latihan guru dan pengalaman mengajar terhadap efikasi guru sekolah menengah.Jurnal Pendidikan Malaysia*, 34 (2). pp. 3-14.

Larson, W. C., & Goebel, A. J. (2008). Putting theory into practice: A professional development school/university co-teaching project. *Journal of the Scholarship of Teaching and Learning*, *8*, 52-61.

Lin, H. & Gorell, J. (2001). Explarotary analysis of pre-service teachers in Taiwan. *Teaching and Teacher Education*, *17*, pp. 623 – 35.

Linacre J.M. (2004) When does a gap between measures matter?, *Rasch Measurement Transactions*, 18:3 p. 993

Messick, S. (1993) Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed.). Washington, D.C.: American Council on Education (Chapter 2)

Luszczynska, A., Scholz, U., & Schwarzer, R. (2005). The general self-efficacy scale: Multicultural validation studies. *Journal of Psychology*, 139, 439-457.

Schwarzer, R., & Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: Mediation analyses. *Applied Psychology, 57*, 152-171.

Smith, C. T. (2008). *An analysis of special education teachers' overall sense of efficacy beliefs and attitudes toward co-taught classrooms*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Tschannen-Moran, M. and Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, pp. 783-805.

Tsigilis, N., Koustelios, A. & Grammatikopoulos, V. (2010). Psychometric properties of the teachers' sense of efficacy scale within the Greek educational context. *Journal of Psychoeducational Assessment, 28(2)*, pp. 153 – 162.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.